# Attachment 15

Article, National Institute of Standards and Technology, "NIST Experts Urge Caution in Use of Courtroom Evidence Presentation Method," October 12, 2017

www.nist.gov/news-events/news/2017/10/nist-experts-urge-caution-use-courtroom-evidence-presentation-method

![NIST - National Institute of Standards and Technology - U.S. Department of Commerce]

# NIST Experts Urge Caution in Use of Courtroom Evidence Presentation Method

## Use of 'Likelihood Ratio' not consistently supported by scientific reasoning approach, authors state.

October 12, 2017



Credit: Hanacek/NIST, isak55-tlegend/Shutterstock

Two experts at the National Institute of Standards and Technology (NIST) are calling into question a method of presenting evidence in courtrooms, arguing that it risks allowing personal preference to creep into expert testimony and potentially distorts evidence for a jury.

The method involves the use of Likelihood Ratio (LR), a statistical tool that gives experts a shorthand way to communicate their assessment of how strongly forensic evidence, such as a fingerprint or DNA sample, can be tied to a suspect. In essence, LR allows a forensics expert to boil down a potentially complicated set of circumstances into a number—providing a pathway for experts to concisely express their conclusions based on a logical and coherent framework. LR's proponents say it is appropriate for courtroom use; some even argue that it is the only appropriate method by which an expert should explain evidence to jurors or attorneys.

However, in a new paper published in the *Journal of Research of the National Institute of Standards and Technology*  (https://doi.org/10.6028/jres.122.027), statisticians Steve Lund and Hari Iyer caution that the justification for using LR in courtrooms is flawed. The justification is founded on a reasoning approach called Bayesian decision theory  (https://en.wikipedia.org/wiki/Bayesian_inference), which has long been used by the scientific community to create logic-based statements of probability. But Lund and Iyer argue that while Bayesian reasoning works well in personal decision making, it breaks down in situations where information must be conveyed from one person to another such as in courtroom testimony.

These findings could contribute to the discussion among forensic scientists regarding LR, which is increasingly used in criminal courts in the U.S. and Europe.

While the NIST authors stop short of stating that LR ought not to be employed whatsoever, they caution that using it as a one-size-fits-all method for describing the weight of evidence risks conclusions being driven more by unsubstantiated assumptions than by actual data. They recommend using LR only in cases where a probability-based model is warranted. Last year's report (https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf) from the President's Council of Advisors on Science and Technology (PCAST) mentions some of these situations, such as the evaluation of high-quality samples of DNA from a single source.

"We are not suggesting that LR should never be used in court, but its envisioned role as the default or exclusive way to transfer information is unjustified," Lund said. "Bayesian theory does not support using an expert's opinion, even when expressed numerically, as a universal weight of evidence. Among different ways of presenting information, it has not been shown that LR is most appropriate."

Bayesian reasoning is a structured way of evaluating and re-evaluating a situation as new evidence comes up. If a child who rarely eats sweets says he did not eat the last piece of blueberry pie, his older sister might initially think it unlikely that he did, but if she spies a bit of blue stain on his shirt, she might adjust that likelihood upward. Applying a rigorous version of this approach to complex forensic evidence allows an

expert to come up with a logic-based numerical LR that makes sense to the expert as an individual.

The trouble arises when other people—such as jurors—are instructed to incorporate the expert's LR into their own decision-making. An expert's judgment often involves complicated statistical techniques that can give different LRs depending on which expert is making the judgment. As a result, one expert's specific LR number can differ substantially from another's.

"Two people can employ Bayesian reasoning correctly and come up with two substantially different answers," Lund said. "Which answer should you believe, if you're a juror?"

In the blueberry pie example, imagine a jury had to rely on expert testimony to determine the probability that the stain came from a specific pie. Two different experts could be completely consistent with Bayesian theory, but one could testify to, say, an LR of 50 and another to an LR of 500—the difference stemming from their own statistical approaches and knowledge bases. But if jurors were to hear 50 rather than 500, it could lead them to make a different ultimate decision.

Viewpoints differ on the appropriateness of using LR in court. Some of these differences stem from the view that jurors primarily need a tool to help them to determine reasonable doubt, not particular degrees of certainty. To Christophe Champod, a professor of forensic science at the University of Lausanne, Switzerland, an argument over LR's statistical purity overlooks what is most important to a jury.

"We're a bit presumptuous as expert witnesses that our testimony matters that much," Champod said. "LR could perhaps be more statistically pure in the grand scheme, but it's not the most significant factor. Transparency is. What matters is telling the jury what the basis of our testimony is, where our data comes from, and why we judge it the way we do."

The NIST authors, however, maintain that for a technique to be broadly applicable, it needs to be based on measurements that can be replicated. In this regard, LR often falls short, according to the authors.

"Our success in forensic science depends on our ability to measure well. The anticipated use of LR in the courtroom treats it like it's a universally observable quantity, no matter who measures it," Lund said. "But it's not a standardized measurement. By its own definition, there is no true LR that can be shared, and the differences between any two individual LRs may be substantial."

The NIST authors do not state that LR is always problematic; it may be suitable in situations where LR assessments from any two people would differ inconsequentially. Their paper offers a framework for making such assessments, including examples for applying them.

Ultimately, the authors contend it is important for experts to be open to other, more suitable science-based approaches rather than using LR indiscriminately. Because

these other methods are still under development, the danger is that the criminal justice system could treat the matter as settled.

"Just because we have a tool, we should not assume it's good enough," Lund said. "We should continue looking for the most effective way to communicate the weight of evidence to a nonexpert audience."

---

Paper: S.P. Lund and H. Iyer, Likelihood Ratio as Weight of Forensic Evidence: A Closer Look. *Journal of Research of National Institute of Standards and Technology*, Published online 12 October 2017. DOI: 10.6028/jres.122.027 (https://doi.org/10.6028/jres.122.027)

# Media Contact

**Chad Boutin** (https://www.nist.gov/people/chad-boutin)

boutin@nist.gov (https://www.nist.govmailto:boutin@nist.gov)
(301) 975-4261

# Related News

- How to Quantify the Weight of Forensic Evidence: A Lively Debate (https://www.nist.gov/news-events/news/2016/06/how-quantify-weight-forensic-evidence-lively-debate)

# Related Links

- Forensic Science (https://www.nist.gov/topics/forensic-science)

*Created October 12, 2017, Updated October 12, 2017*